

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 16, 2003, 16:55:57 ; Search time 440.214 Seconds  
(without alignments)  
330.553 Million cell updates/sec

Title: US-09-856-070-16  
 Perfect score: 25  
 Sequence: 1 EREKE 5

Scoring table:	BL0SUM62		
	Xgapop 10.0	Xgapext 0.5	
	Ygapop 10.0	Ygapext 0.5	
	Egapop 6.0	Egapext 7.0	
	Delop 6.0	Delext 7.0	

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 416,280

```
Minimum DB seq length: 0
Maximum DB seq length: 200000000
```

Post processing. Minimum Match: 0%  
Maximum Match: 100%  
Listing first 45 summaries

```

Command line parameters:
-MemDir=ram0_128M.spec; -F2M.spec; -DEV-xip
-o /cygn2/128M.spec; -libdir=/usr/lib64; -gnat-1401200; -155934; -1631; -app-queue; fasta_1_1502
-MemDir=ram0_128M.spec; -F2M.spec; -DEV-xip
-o /cygn2/128M.spec; -libdir=/usr/lib64; -gnat-1401200; -155934; -1631; -app-queue; fasta_1_1502
-DB-Genbank1 -Qfmt-fastap -Suffix-idx -MINMATCH=0.1 -LocuP=0.0 -DB-EXE=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human4.0 -cd1 -LIST=45
-DOWPARAM=200 -THER-SCORE=1 -THER-MAX=100 -THER-MIN=0 -ALLIN=15 -MODE=LOCAL
-COMMIT=100 -NUM-CORE= -HEAPSIZE=500 -MINI=0 -MAXLEN=2000000000
-USEP=7509856070 -argn_1_15741 -argu1_1401200; -155934; -1631; -NTH=4; -LTH=3
-NO_XLIP; NO_MMAPP -LARGEXE=QUERY -NE5 SIZE=0.0 -WALL -UNTH=0 -DEV -TIMO=0.120
-WARN_TIME=0.10 -TUPEDAS=1 -XMAP=0.10 -XCAPEXT=0.5 -PCAP=0.6 -FCAPEXT 7
-VCAPOP=10 -VCAPEXT=0.5 -DEIOP=10 -XCAPEXT=0.5 -PCAP=0.6 -FCAPEXT 7

```

Database : GenBank : ★

1: qb\_ba:\*  
2: qb\_bt:\*  
3: qb\_in:\*  
4: qb\_cm:\*  
5: qb\_ov:\*  
6: qb\_pr:\*  
7: qb\_ph:\*  
8: qb\_pl:\*  
9: qb\_pr:\*  
10: qb\_ro:\*  
11: qb\_sis:\*  
12: qb\_sy:\*  
13: qb\_un:\*  
14: qb\_vl:\*  
15: em\_ba:\*  
16: em\_lun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_cm:\*  
21: em\_of:\*  
22: em\_pa:\*  
23: em\_pa:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_tr:\*  
27: em\_un:\*  
28: em\_un:\*

```

29: em_vi: *
30: em_htq_hum: *
31: em_htq_irv: *
32: em_htq_other: *
33: em_htq_mus: *
34: em_htq_pln: *
35: em_htq_rtd: *
36: em_htq_mam: *
37: em_htq_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query			ID	Description
		Match	Length	PH		
C 1	25	100.0	42	6	AF089429	Sequence
C 2	25	100.0	22	6	AR158374	Sequence
C 3	25	100.0	22	6	AR172577	Sequence
C 4	25	100.0	25	6	AF137561	Sequence
C 5	25	100.0	25	6	AF139067	Sequence
C 6	25	100.0	47	6	AX378317	Sequence
C 7	25	100.0	58	6	AR208349	Sequence
C 8	25	100.0	60	8	S44200	class V1.2y
C 9	25	100.0	61	8	AF062766	Sequence
C 10	25	100.0	62	4	AF164071	Sequence
C 11	25	100.0	64	11	HUMU1272A	Sequence
C 12	25	100.0	65	6	AX483825	Sequence
C 13	25	100.0	67	10	MCMNC52b	Sequence
C 14	25	100.0	74	7	PF1TP3	Sequence
C 15	25	100.0	84	8	AX440980	Sequence
C 16	25	100.0	84	8	ATH243705	Sequence
C 17	25	100.0	84	10	AF200762S4	Sequence
C 18	25	100.0	87	5	AF033554	Sequence
C 19	25	100.0	91	10	MMD1RE31	Sequence
C 20	25	100.0	95	11	HMWTS301A	Sequence
C 21	25	100.0	96	3	AF411699	Sequence
C 22	25	100.0	100	3	AF411993	Sequence
C 23	25	100.0	102	4	AF317797	Sequence
C 24	25	100.0	103	8	AY074729	Sequence
C 25	25	100.0	110	4	AF317798	Sequence
C 26	25	100.0	110	4	AF317799	Sequence
C 27	25	100.0	110	4	AF317800	Sequence
C 28	25	100.0	111	4	AF155792	Sequence
C 29	25	100.0	111	9	HUMHE1AF	Sequence
C 30	25	100.0	112	4	AF317793	Sequence
C 31	25	100.0	112	4	AF317794	Sequence
C 32	25	100.0	112	4	AF317795	Sequence
C 33	25	100.0	112	4	AF317796	Sequence
C 34	25	100.0	113	11	AJ028712	Sequence
C 35	25	100.0	115	3	ACU420073	Sequence
C 36	25	100.0	115	3	HS15HG6F	Sequence
C 37	25	100.0	116	4	CFEAL037	Sequence
C 38	25	100.0	119	5	AF114759	Sequence
C 39	25	100.0	119	10	AF250644	Sequence
C 40	25	100.0	120	3	AF250645	Sequence
C 41	25	100.0	120	9	Y025293811	Sequence
C 42	25	100.0	123	11	GA4924	Sequence
C 43	25	100.0	123	11	GA4925	Sequence
C 44	25	100.0	123	11	GE3882	Sequence
C 45	25	100.0	125	11	GG2146	Sequence

## ALLIANCE PARTNERS

## RESULT 1

AK089240  
LOCUS AK089240  
DEFINITION Sequence 19 from patent US 5994062.  
ACCESSION AK089240  
VERSION AK089240.1 GI:10015487  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 22)  
AUTHORS Mulshine, J.L. and Lockman, M.S.  
TITLE Epithelial protein and DNA thereof for use in early cancer detection  
JOURNAL Patent: US 5994062-A 19 30-NOV-1999;  
FEATURES Location/Qualifiers  
source 1..22  
BASE COUNT 10 a 3 c 7 q 2 t  
ORIGIN /organism="unknown"

Alignment Scores:  
Pred. No.: 31.6 Length: 22  
Score: 25.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-856-070-16 (1-5) x AK089240 (1-22)

QY 1 GluArgGluLysGlu 5  
Db 1 GACAGAGAAAAGGAA 15

RESULT 2  
LOCUS AK158374  
DEFINITION Sequence 19 from patent US 6251586.  
ACCESSION AK158374  
VERSION AK158374.1 GI:16220396  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 22)  
AUTHORS Mulshine, J.L. and Lockman, M.S.  
TITLE Epithelial protein and DNA thereof for use in early cancer detection  
JOURNAL Patent: US 6251586-A 19 26-JUN-2001;  
FEATURES Location/Qualifiers  
source 1..22  
BASE COUNT 10 a 3 c 7 q 2 t  
ORIGIN /organism="unknown"

Alignment Scores:  
Pred. No.: 31.6 Length: 22  
Score: 25.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-856-070-16 (1-5) x AK158374 (1-22)

QY 1 GluArgGluLysGlu 5  
Db 1 GACAGAGAAAAGGAA 15

RESULT 4  
LOCUS AK172577  
DEFINITION Sequence 47 from patent US 6197517.  
ACCESSION AK172577  
VERSION AK172577.1 GI:14474070  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 25)  
AUTHORS Roberts, C.J.  
TITLE Essential genes of yeast as targets for antifungal agents, herbicides, insecticides and anti-proliferative drugs  
JOURNAL Patent: US 6197517-A 47 06-MAR-2001;  
FEATURES Location/Qualifiers  
source 1..25  
BASE COUNT 4 a 8 c 1 q 12 t  
ORIGIN /organism="unknown"

Alignment Scores:  
Pred. No.: 36.1 Length: 25  
Score: 25.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-856-070-16 (1-5) x AK172577 (1-25)

QY 1 GluArgGluLysGlu 5  
Db 19 GACAGAGAAAAGGAG 5

RESULT 5  
LOCUS AK139067  
DEFINITION Sequence 47 from patent US 6200803.  
ACCESSION AK139067  
VERSION AK139067.1 GI:14481412  
KEYWORDS

ACCESSION AK172577  
VERSION AK172577.1 GI:14474070  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 22)  
AUTHORS Re, R. and Cook, J.  
TITLE Inhibition of cellular proliferation in vitro by oligonucleotide binding to a chromosomal binding site for p53 protein  
JOURNAL Patent: US 6303328-A 9 16-OCT-2001;  
FEATURES Location/Qualifiers  
source 1..22  
BASE COUNT 0 a 7 c 0 q 15 t  
ORIGIN /organism="unknown"

Alignment Scores:  
Pred. No.: 31.6 Length: 22  
Score: 25.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-856-070-16 (1-5) x AK172577 (1-22)

QY 1 GluArgGluLysGlu 5  
Db 22 GAAAGAGAAAAGGAA 8

RESULT 4  
LOCUS AK137561  
DEFINITION Sequence 47 from patent US 6197517.  
ACCESSION AK137561  
VERSION AK137561.1 GI:14474070  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 25)  
AUTHORS Roberts, C.J.  
TITLE Essential genes of yeast as targets for antifungal agents, herbicides, insecticides and anti-proliferative drugs  
JOURNAL Patent: US 6197517-A 47 06-MAR-2001;  
FEATURES Location/Qualifiers  
source 1..25  
BASE COUNT 4 a 8 c 1 q 12 t  
ORIGIN /organism="unknown"

Alignment Scores:  
Pred. No.: 36.1 Length: 25  
Score: 25.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-856-070-16 (1-5) x AK137561 (1-25)

QY 1 GluArgGluLysGlu 5  
Db 19 GACAGAGAAAAGGAG 5

RESULT 5  
LOCUS AK139067  
DEFINITION Sequence 47 from patent US 6200803.  
ACCESSION AK139067  
VERSION AK139067.1 GI:14481412  
KEYWORDS

SOURCE: Unknown.  
ORGANISM: Unknown.  
REFERENCE: 1 (bases 1 to 25)  
AUTHORS: Roberts, C.J.  
TITLE: Essential genes of yeast as targets for antifungal agents, herbicides, insecticides and anti-proliferative drugs  
JOURNAL: Patent: US 620803-A 47 13-MAR-2001.  
FEATURES: Location/Qualifiers  
SOURCE: 1..25  
BASE COUNT: 4 a 8 c 1 g 12 t  
ORIGIN: 1..25  
Alignment Scores:  
Pred. No.: 36.1 Length: 25  
Score: 25.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-856-070-16 (1-5) x APL30067 (1-25)  
QY 1 GluArgGluLysGlu 5  
Db 19 GAGAGAGAGAGAGAG 5  
RESULT 6  
LOCUS AX378317/c 47 bp DNA linear PAT 18-MAR-2002  
DEFINITION Sequence 106 from Patent WO0206525.  
ACCESSION AX378317  
VERSION AX378317.1 GI:19574167  
KEYWORDS human.  
SOURCE: Homo sapiens  
ORGANISM: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates, Catarrhini; Hominoidea; Homo.  
REFERENCE 1  
AUTHORS Cohen, D., Blumenfeld, M., Chumakov, I., Abderrahim, H. and Bihaoui, R.  
TITLE Obesity associated biallelic marker maps  
JOURNAL Patent: WO 0206525-A 106 24-JAN-2002;  
GENSET (FR)  
FEATURES: Location/Qualifiers  
SOURCE: 1..47  
variation /organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT: 8 a 13 c 2 g 23 t 1 others  
ORIGIN: 1..47  
Alignment Scores:  
Pred. No.: 69.8 Length: 47  
Score: 25.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-856-070-16 (1-5) x AX378317 (1-47)  
QY 1 GluArgGluLysGlu 5  
Db 18 GAAAGAGAGAGAGAG 4  
RESULT 7  
LOCUS AR208349/c 58 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 5 from patent US 6383747.  
ACCESSION AR208349  
VERSION AR208349.1 GI:21509480

KEYWORDS: Unknown.  
SOURCE: Unknown.  
ORGANISM: Unknown.  
REFERENCE: 1 (bases 1 to 58)  
AUTHORS Dawkins, R. Letts, and Abraham, L. Joseph.  
TITLE Method for determining ancestral haplotypes using haplo-specific genetic elements within the major histocompatibility complex multigene cluster  
JOURNAL: Patent: US 648747-A 5 07-MAY-2002.  
FEATURES: Location/Qualifiers  
SOURCE: 1..58  
BASE COUNT: 1 a 27 c 0 g 30 t  
ORIGIN: 1..58  
Alignment Scores:  
Pred. No.: 87 Length: 58  
Score: 25.00 Matches: 5  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-856-070-16 (1-5) x AR208349 (1-58)  
QY 1 GluArgGluLysGlu 5  
Db 50 GAGAGAGAGAGAGAG 36  
RESULT 8  
LOCUS S44200/c 60 bp mRNA linear PLN 08-MAY-1993  
DEFINITION class VI zygote specific cell wall hydroxyproline-rich glycoprotein [Chlamydomonas reinhardtii, mRNA Partial, 60 nt].  
ACCESSION S44200  
VERSION S44200.1 GI:255403  
KEYWORDS: Chlamydomonas reinhardtii.  
ORGANISM: Chlamydomonas reinhardtii  
REFERENCE 1 (bases 1 to 60)  
AUTHORS Woessner, J. P. and Goodenough, U. W.  
TITLE Molecular characterization of a zygote wall protein: an extensin-like molecule in Chlamydomonas reinhardtii  
JOURNAL Plant Cell 1 (9), 901-911 (1989)  
MEDLINE 92393420  
PubMed 2535530  
REMARK GenBank Staff at the National Library of Medicine created this entry [NCBI gisbq 113226] from the original journal article. This sequence comes from Fig 6  
FEATURES: Location/Qualifiers  
SOURCE: 1..60  
/organism="Chlamydomonas reinhardtii"  
/db\_xref="taxon:3055"  
gene 1..60  
/partial  
/gene="class VI zygote specific cell wall hydroxyproline-rich glycoprotein"  
CDS 1..60  
/partial  
/gene="class VI zygote specific cell wall hydroxyproline-rich glycoprotein"  
/note="This sequence comes from Fig. 6"  
/codon\_start=1  
/protein\_id="AA023259.1"  
/db\_xref="GI:255404"  
BASE COUNT: 3 a 35 c 5 g 17 t  
ORIGIN: 3 a 35 c 5 g 17 t  
Alignment Scores:



\_\_\_\_\_

Alignment Scores:  
 Pred. No.: 119 Length: 78  
 Score: 25.00 Matches: 5  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 7 Gaps: 0

US 09-856-070-16 (1-5) x PFITR4 (1-78)

QY 1 GluArgGlnLysGlu 5  
 |||||  
 Db 10 GAGACAGAGAAAGAA 24

# RESULT 15

AX240980/c

LOCUS AX240980 24 bp DNA linear PAT 26-SEP-2001

DEFINITION Sequence 218 from Patent WO0160975.

ACCESSION AX240980

VERSION AX240980.1 GI:15797855

## KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct.

artificial sequences.

REFERENCE 1 (bases 1 to 83)

AUTHORS Roemer, T., Jiang, B., Hoone, C. and Hussey, H.

TITLE Gene disruption methodologies for drug target discovery

JOURNAL Patent: WO 0160975-A 218 23-AUG-2001;

Elitra Pharmaceuticals, Inc. (US)

## FEATURES

source location/Qualifiers

1..83  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="DNA primer"

BASE COUNT 21 a 14 c 11 g 37 t

## ORIGIN

Alignment Scores:  
 Pred. No.: 127 Length: 83  
 Score: 25.00 Matches: 5  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US 09-856-070-16 (1-5) x AX240980 (1-83)

QY 1 GluArgGlnLysGlu 5  
 |||||  
 Db 34 GAGACAGAGAAAGAG 20

Search completed: January 16, 2003, 19:02:22  
 Job time : 444,214 secs